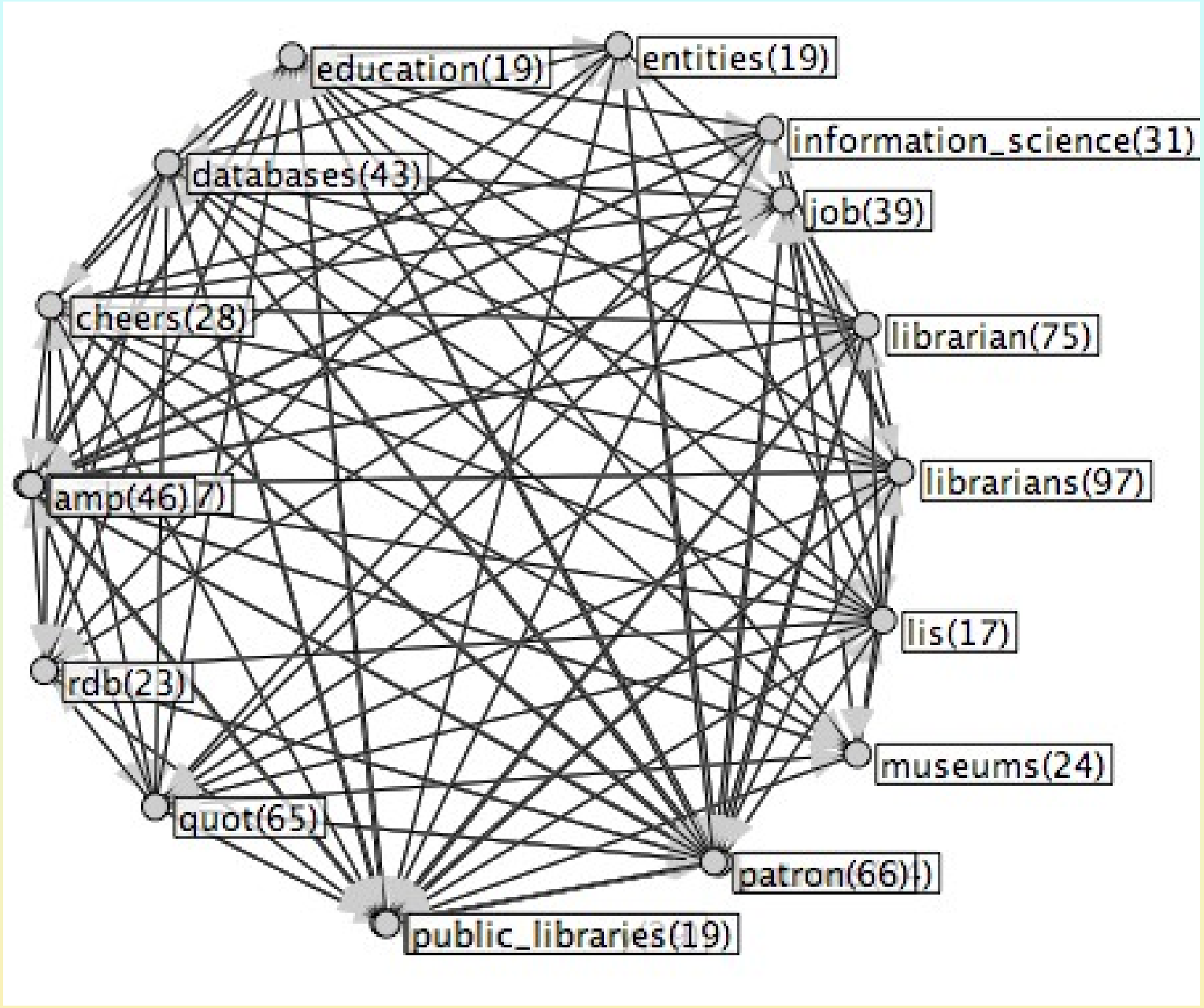

education(19)
entities(19)
information_science(31)
databases(43)
job(39)
cheers(28)
librarian(75)
amp(46)
librarians(97)
lis(17)
rdb(23)
museums(24)
quot(65)
patron(66)
public_libraries(19)

Lectures at IBICT, June 2009

- These powerpoint slides accompanied one of series of lectures given in June 2009 at IBICT, Instituto Brasileiro de Informação em Ciência e Tecnologia (The Brazilian Institute for Information in Science and Technology), Rio de Janeiro, where Professor Caroline Haythornthwaite was a guest of the institute.
- Thanks go to
  - Celia Ribeiro Zaher, Coordenadora de Ensino e Pesquisa, C&T da Informação, IBICT for arranging this visit.
  - Professora Gilda Olinto for working with me on this series and translating my slides into Portuguese
- Lectures included: An overview of e-learning; Computer-mediated communication (CMC) and e-learning; Social informatics (SI) and e-learning; E-learning networks; Theories and ideas emerging for e-learning; Networked learning

Analisando textos de E-learning

- Provide operational platform for visualization of interaction and group dynamics in online conversations
  - Recapture the visibility of the ‘group’ in online communication
  - Augment linear text-based representation with spatial representation
  - For feedback for participants, and overviews for instructors
- Provide analytical platform for comparison across cases
  - Define concepts and metrics useful for analysis of group or community interaction
  - Identify interactional relations and patterns that are key for successful online learning practice and experience

The Stimulus

- Rios de textos lineares
- Estrutura social invisível
- Rapidamente gerada
- Padrões de participação invísíveis

March 2008 Archives by thread
. Messages sorted by: [subject] [author] [date]
. More info on this list...
Starting: Sat Mar 1 13:26:33 PST 2008_Ending: Sun Mar 23 09:32:23 PDT 2008_Messages: 205

. [Air-L] Final goodbye for early web icon Dominic Pinto
. [Air-L] Reminder - Gogimon Search Agent Beta Tesers David Miller
. [Air-L] Meeting in Illinois, May 08 - "decolonized methodologies" Denise N. Rall
. [Air-L] Open Source and changing mode of productions in the third world Denise N. Rall
. [Air-L] Companion to Digital Humanities Barry Wellman
. [Air-L] Companion to Digital Humanities Jankowski
. [Air-L] Origins of E-Commerce Alex -Vipowernet
. [Air-L] Lessons in Second LIfe jeremy hunsinger
. [Air-L] Invitation to 6th Annual Workshop on Open and User Innovation - HBS & MIT - August 4-6, 2008 Karim R. Lakhani
. [Air-L] Facing up to Facebook - Michael Geist at Osgoode March 5 (Livecast available!) Giuseppina D'Agostino/osgoode
. [Air-L] Instrument help: eveluate user's perception of onl ine community Ke, Nan
. [Air-L] CFP - DIAC Demos, Workshops, and Exploratory Papers Tom Erickson
. [Air-L] call for papers for a special issue on consumption and Web 2.0 davidgbeer at aol.com
. [Air-L] TVO The Agenda tonight Nancy Baym
. [Air-L] Resources on On-Line Dating and SMS Language Andrew Herman
. [Air-L] Resources on On-Line Dating and SMS Language Gordon Carlson
. [Air-L] Postdoc in new media (Germany/Switzerland) Elad Segev
. [Air-L] Online research ethics Alecea Standlee
. [Air-L] Online research ethics Nishant Shah
. [Air-L] Online research ethics mhward
. [Air-L] Online research ethics joana ro
. [Air-L] Online research ethics Charles Ess
. [Air-L] Online research ethics - my two and 1/4 cents Radhika Gajjala
. [Air-L] Online research ethics - my two and 1/4 cents Lois Ann Scheidt
. [Air-L] Online research ethics Jim Porter
. [Air-L] Online research ethics Radhika Gajjala
. [Air-L] Online research ethics coopman at u.washington.edu
. [Air-L] Online research ethi cs Charles Ess
. [Air-L] Online research ethics Jeremy Hunsinger
. [Air-L] Online research ethics Jeremy Hunsinger
. [Air-L] Online research ethics Heidelberg, Chris
. [Air-L] Online research ethics Derek Hansen
. [Air-L] Online research ethics Jeremy Hunsinger
. [Air-L] Online research ethics Steve Jones
. [Air-L] Online research ethics Lois Ann Scheidt
. [Air-L] Online research ethics Radhika Gajjala

. [Air-L] Online research ethics Jeremy Hunsinger
. [Air-L] Online research ethics Charlie Balch
. [Air-L] Online research ethics Lois Ann Scheidt
. [Air-L] Online research ethics Jeremy Hunsinger
. [Air-L] Online research ethics Steve Jones
. [Air-L] Online research ethics Andrew Rojecki
. [Air-L] Online research ethics Mark D. Johns
. [Air-L] Online research ethics - SL Radhika Gajjala
. [Air-L] Online research ethics - SL Lois Ann Scheidt
. [Air-L] Online research ethics - SL Radhika Gajjala
. [Air-L] avatar research ethics Jeremy Hunsinger
. [Air-L] Online research ethics Mark D. Johns
. [Air-L] Postdoc in new media (Germany/Switzerland) Geder Parzianello
. [Air-L] FW: 'Digital Ontario' A symposium Wednesday, March 5th, 2008 - Thursday, March 6th, 2008 Michael Gurstein
. [Air-L] CFP: HICSS 42 : Social Networks and Virtual Worlds for Work, Learning, and Play Caroline Haythornthwaite
. [Air-L] Invitation to Participate: Research Related to Internet Governance Nanette Levinson
. [Air-L] IP/Gender 4/4/08 burkx006 at umn.edu
. [Air-L] Top ten web apps mhward
. [Air-L] Online research ethics Jankowski
. [Air-L] REMINDER> 15 March deadline for e-Research 08 conference in Oxford Eric T. Meyer
. [Air-L] Politics: Web 2.0, Royal Holloway, University of London LAST CHANCE TO REGISTER Chadwick Andrew
. [Air-L] Online research ethics Marj Kibby
. [Air-L] avatar research ethics Marj Kibby
. [Air-L] avatar research ethics Radhika Gajjala
. [Air-L] avatar research ethics Steve Jones
. [Air-L] avatar research ethics Gordon Carlson
. [Air-L] avatar research ethics Radhika Gajjala
. [Air-L] avatar research ethics Kristin Lindsley
. [Air-L] avatar research ethics Gordon Carlson
. [Air-L] TorontoStar: Facebook : The New Study Hall For The Wired Generation? Perhaps Not [:-(Gerry Mckiernan
. [Air-L] TorontoStar: Facebook : The New Study Hall For The Wired Generation? Perhaps Not [:-(Greg Elmer
. [Air-L] TorontoStar: Facebook : The New Study Hall For TheWired Generation? Perhaps Not [:-(Marj Kibby
. [Air-L] TorontoStar: Facebook : The New Study Hall For TheWired Generation? Perhaps Not [:-(Dr. Steve Eskow
. [Air-L] TorontoStar: Facebook : The New Study Hall For TheWired Generation? Perhaps Not [:-(Peter Timusk
. [Air-L] TorontoStar: Facebook : The New Study Hall For TheWired Generation? Perhaps Not [:-(Peter Timusk
. [Air-L] Online research ethics dddumitr at ucalgary.ca
. [Air-L] Online research ethics Radhika Gajjala

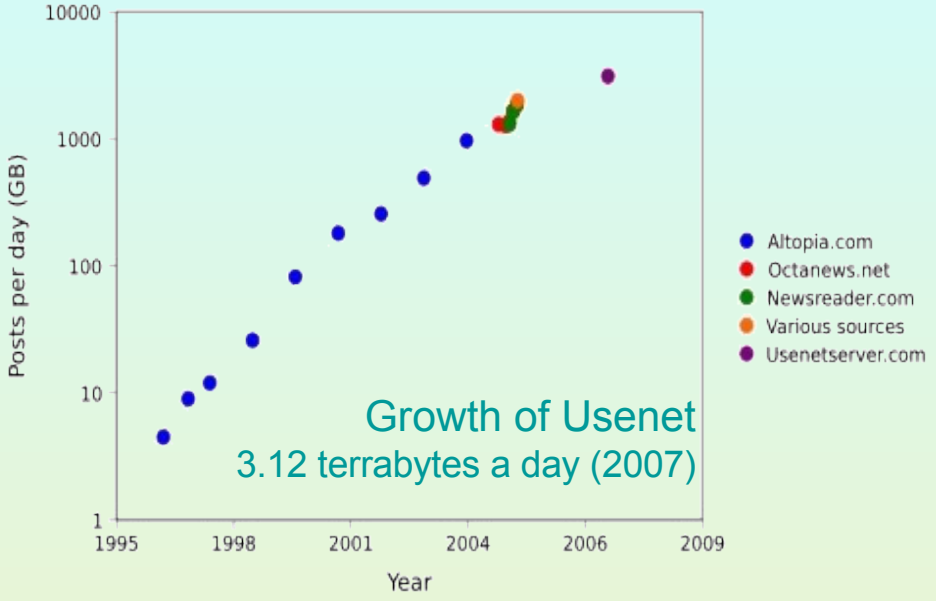
The Stimulus

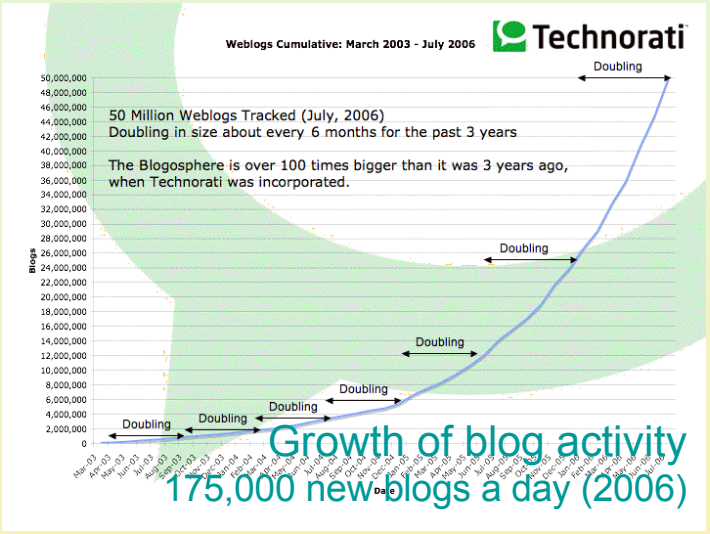
- Growing volume of texts contributed by a growing number of participants
  - Increased amount of online text
  - Increased use of online environments
  - Greater need to understand online interaction processes

" Growth of online learning

  Almost 3.5 million students (US) were taking at least one online course during the fall 2006 term; (10% increase over 2005)

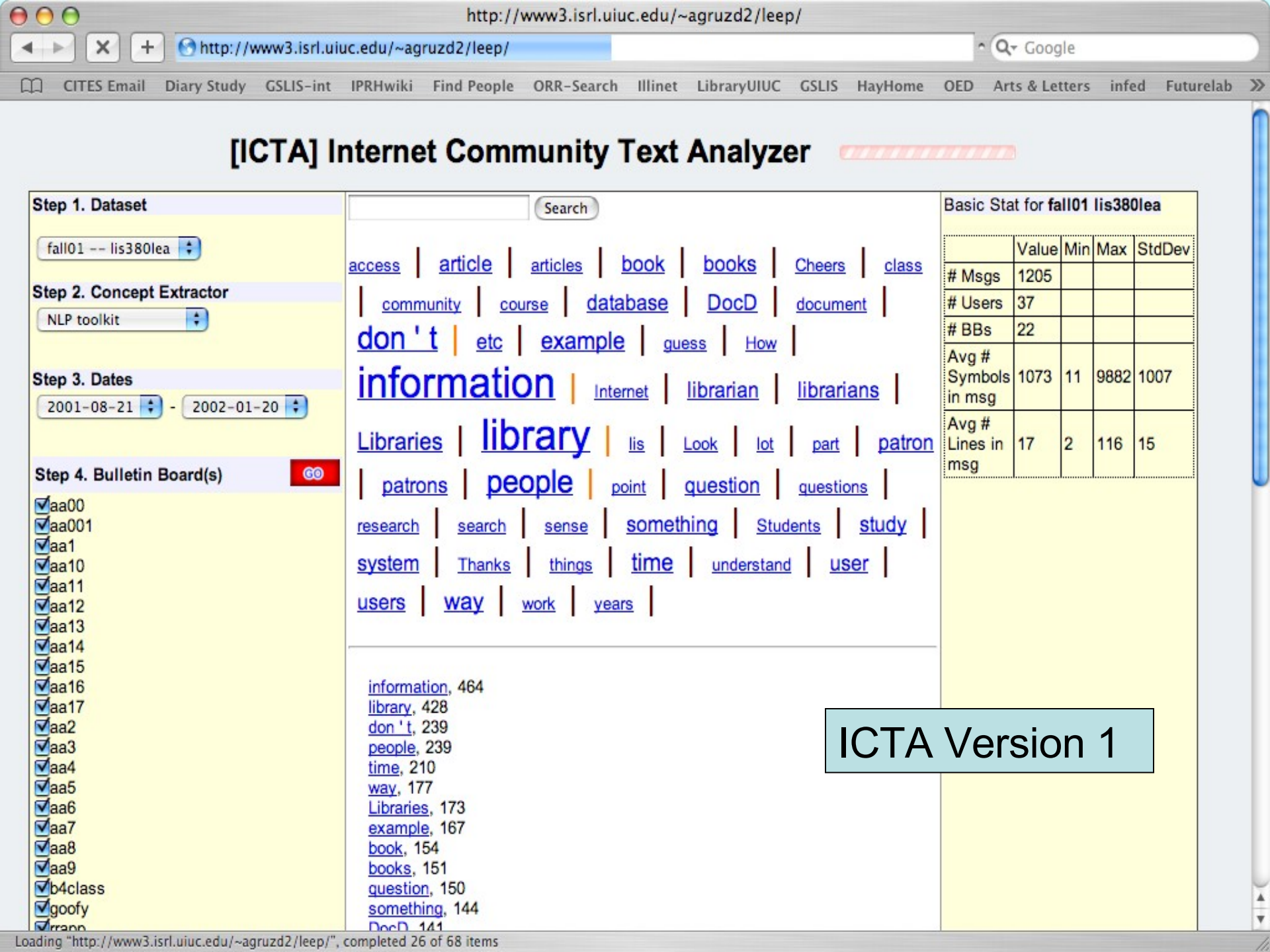
  ~20% of all U.S. higher education students were taking at least one online course in the fall of 2006 (Allen & Seaman, 2007)

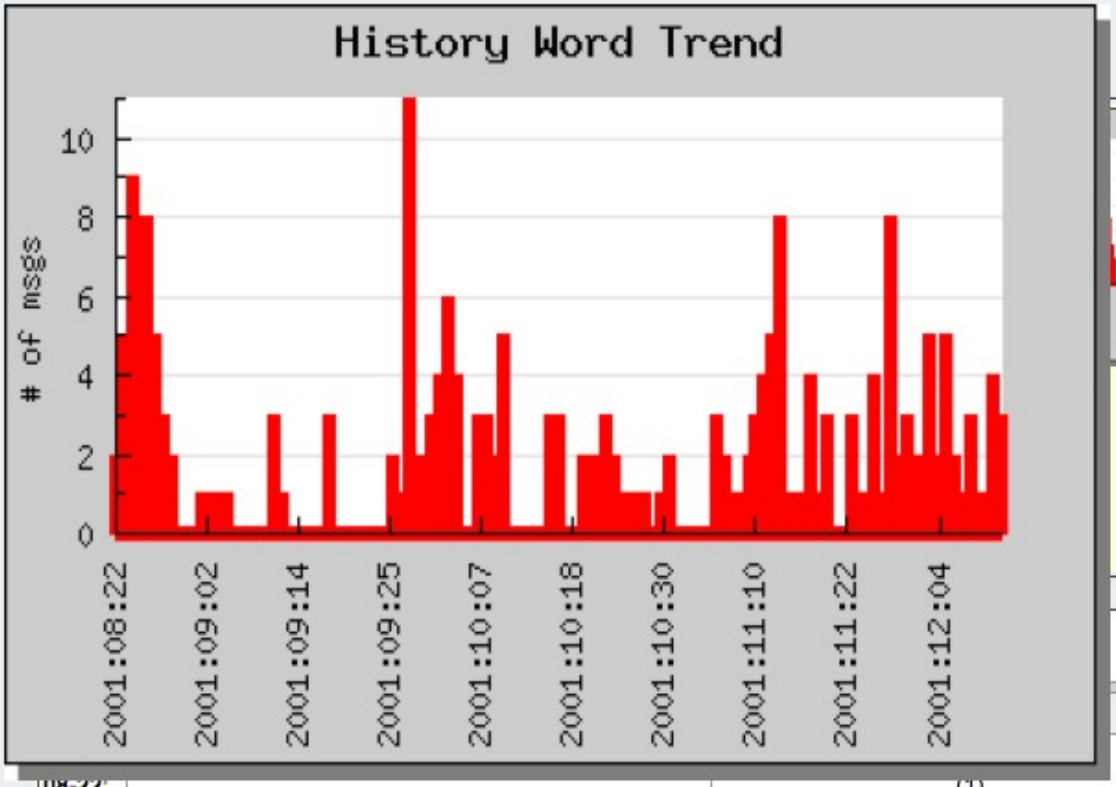




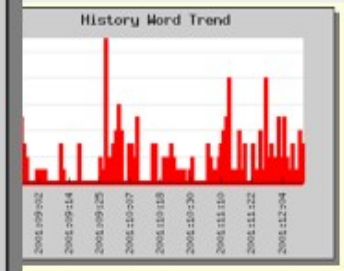
Buscando sentido da ação conversasional

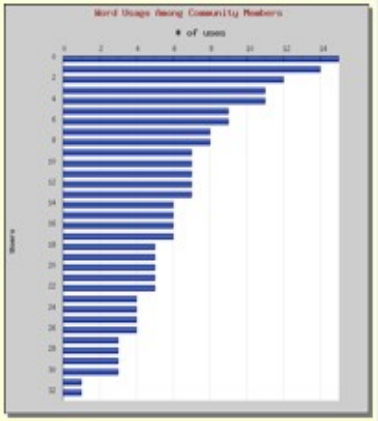
- Usando abordagem de redes sociais (SNA)
- Enfatizando a descoberta empírica da base relacional da interação social e o exame da estrutura social
- Usando processamento de linguagem natural (NLP)
  - Extrair tópicos chave da conversação
- Combinar SNA & NLP
  - Estrair mais nuances
  - Apresentação de textos mais integrada
- Prover visualizações
  - Interatividade, dinâmica de grupo, uso da linguagem


http://www3.isrl.uiuc.edu/~agruzd2/leep/
http://www3.isrl.uiuc.edu/~agruzd2/leep/
Google
CITES Email
Diary Study
GSLIS-int
IPRHwiki
Find People
ORR-Search
Illinet
LibraryUIUC
GSLIS
HayHome
OED
Arts & Letters
infed
Futurelab
[ICTA] Internet Community Text Analyzer
Step 1. Dataset
fall01 -- lis380lea
Step 2. Concept Extractor
NLP toolkit
Step 3. Dates
2001-08-21
2002-01-20
Step 4. Bulletin Board(s)
GO
aa00
aa001
aa1
aa10
aa11
aa12
aa13
aa14
aa15
aa16
aa17
aa2
aa3
aa4
aa5
aa6
aa7
aa8
aa9
b4class
goofy
Search
access | article | articles | book | books | Cheers | class | community | course | database | DocD | document | don ' t | etc | example | guess | How | information | Internet | librarian | librarians | Libraries | library | lis | Look | lot | part | patron | patrons | people | point | question | questions | research | search | sense | something | Students | study | system | Thanks | things | time | understand | user | users | way | work | years
information, 464
library, 428
don ' t, 239
people, 239
time, 210
way, 177
Libraries, 173
example, 167
book, 154
books, 151
question, 150
something, 144
Basic Stat for fall01 lis380lea
Value Min Max StdDev
Msgs 1205
Users 37
BBs 22
Avg # Symbols in msg 1073 11 9882 1007
Avg # Lines in msg 17 2 116 15
ICTA Version 1
Loading "http://www3.isrl.uiuc.edu/~agruzd2/leep/", completed 26 of 68 items



close





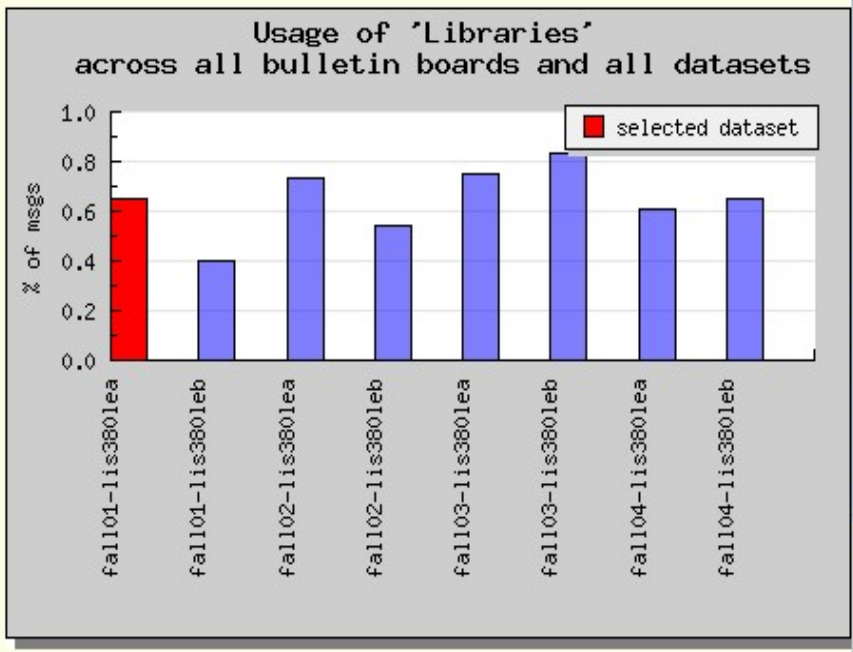
			to ultimately fill these information gaps... a bi
08-22:		(1)	as well. Even though Benjamin D'Isreali said
2001-08-23:	t you mean XXX. When the article talked about how	Libraries (1)	should have access to that type of information, I
2001-08-23:	e for the database would distribute it annually to	Libraries (1)	and other social service agencies on a CD. Think
2001-08-23:	Not one person thought of their public library!	Libraries (2)	still seemed, in 1991 at least when this study wa
2001-	You bring up a good point relating to the role of	Libraries	a decade after this study was done. I would asse

http://www3.isrl.uiuc.edu/~agruzd2/leep/

Google

CITES Email | Diary Study | GSLIS-int | IPRHwiki | Find People | ORR-Search | Illinet | LibraryUIUC | GSLIS | HayHome | OED | Arts & Letters | infed | Futurelab

[ICTA] Internet Communit



Step 1. Dataset

fall01 -- lis380lea

Step 2. Concept Extractor

NLP toolkit

Step 3. Dates

2001-08-21 - 2002-01-20

Step 4. Bulletin Board(s) GO

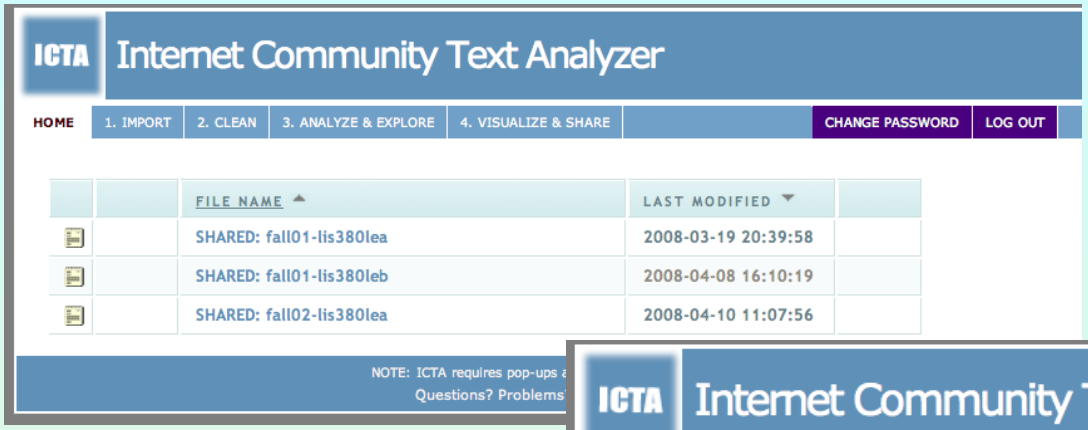
- aa00
- aa001
- aa1
- aa10
- aa11
- aa12
- aa13
- aa14
- aa15
- aa16
- aa17
- aa2
- aa3
- aa4
- aa5
- aa6
- aa7
- aa8
- aa9
- b4class
- goofy

Search

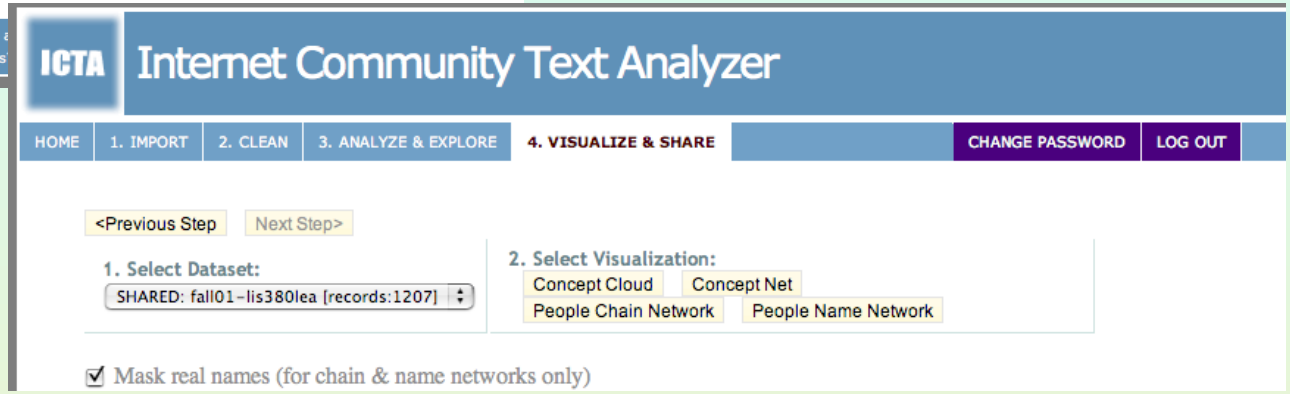
access | article | articles | book | community | course | database | don ' t | etc | example | information | Internet | Libraries | library | lis | patrons | people | point | research | search | sense | system | Thanks | things | time | users | way | work | years

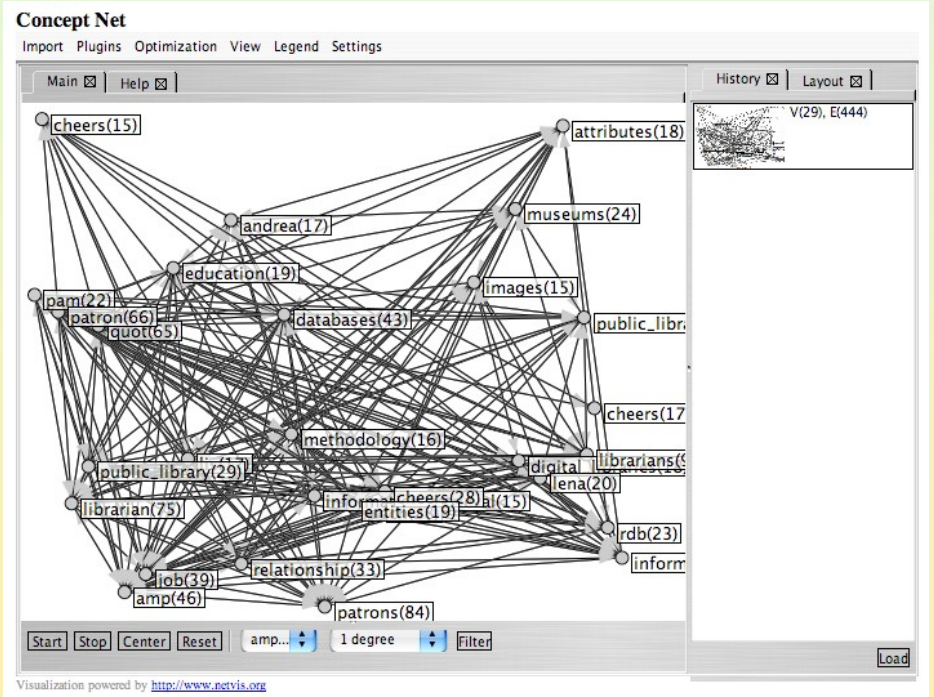
information, 464
library, 428
don ' t, 239
people, 239
time, 210
way, 177
Libraries, 173
example, 167
book, 154
books, 151
question, 150

Run script "%20Mouseclick_txt('get_noun_pairs.php?NNextractorID=nn&noun=Libraries&sql_where='+get_sqlwhere());"



www.textanalytics.com





ICTA version 2

- Concept clouds and networks
- New network construction techniques

Opções de construção de redes

- O foco presente é na descoberta da rede
  - **Chain Network (Rede em cadeia)**
    - Simple, based on who posted and the order of postings, no examination of text
  - **Subject Line** Text Chain **Networks (Rede de linha de assuntos)**
    - Middle complexity, based comparing the text in subject lines, and then using who posts after whom with the same subject line
  - Message Text **Name Network (Rede de nomes)**
    - More complex, based on analyzing the text of messages for names used in the body of postings

Descoberta de Rede

Ex.1

Previous post is by Gabriel, Sam replies:
*'**Nick, Ann, Gina, Gabriel***:
I apologize for not backing this up with a good source,
but I know from reading about this topic that libraries…'

Previous posts by Gabriel, Sam, Gina, and Eva, then:
'Gina, I owe you a cookie. This is exactly what I wanted to know.
I was already planning on taking 302 next semester,
and now I have something to look forward to!'

Ex.2

Ex.3

Post by Fred:
'I wonder if that could be why other libraries
around the world have resisted changing –
it's too much work, and as ***Dan*** pointed out, too expensive.'

Construindo redes a partir de textos de mensagens

- Uso de informação de nós e laços que está no texto das mensagens
- Questões
  - Discovering names and nicknames used in the text
  - Identifying names of people in the class from names of authors being discussed
  - Identifying all the names one person might use (e.g., James, Jim, jw@illinois.edu, jwill@gmail.com)
  - Distinguishing between two or more people with the same name (Jim G. and Jim M.)

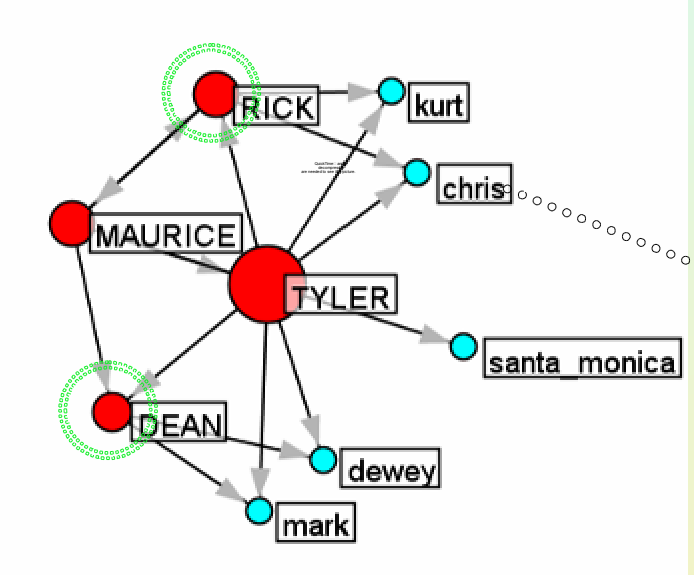
Identificação do ator

- Descoberta de nomes pessoais
  - Class lists of names don't always work
    - e.g., if someone uses their middle name which is not on the name list, or they use a short or nickname
- Método
  - The 1990 US Census http://www.census.gov/genealogy/names
    - Limitation at present is the emphasis on US names
  - Capitalization
  - Context words
    - "Hi, Sammy"
    - "Good night, Jill"

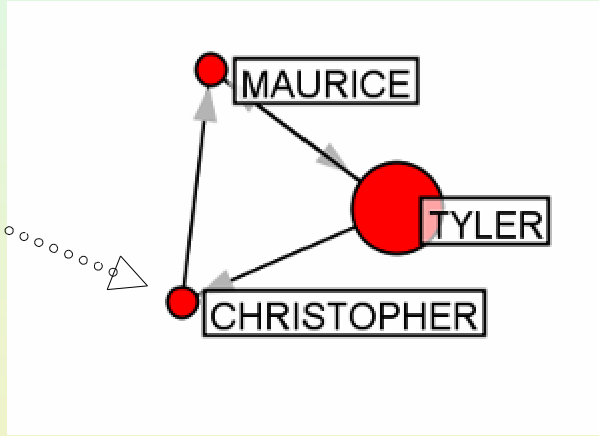
Examplo 2: Extraindo os Nomes

Egocentric network for “Tyler”

Name Network



Chain Network



Visualization powered by http://www.netvis.org

Example 2: Extraindo os Nomes

EXAMPLE
From: wilma@bedrock.us (= **Wilma**)
Reference Chain: tank123@gl.edu, hle@gl.edu

Hi **Dustin**, **Sam** and all, I appreciate your posts from this and last week […]. I keep thinking of poor **Charlie** who only wanted information on "dogs". […] Cheers, **Wilma**.

Words to the Left	Name	Words to the Right	Position	Context word?	FROM	TO
Hi	**Dustin**	Sam and	0	Yes	0	1 · 2 = **2**
HiDustin	**Sam**	and all	0.01	Yes	0.01	0.99 · 2 = **1.98**
Ofpoor	**Charlie**	who only	0.50	No	0.50	0.50
Cheers	**Wilma**		0.88	Yes	0.88 · 2 · 2 = **3.52**	0.12

Algorithm calculates weights for the likelihood that a name is associated with a sender or receiver. Weight under FROM points to Wilma as the poster.
Weight under TO indicates Dustin and Sam as receivers.

Comparando redes de cadeias e redes de nomes

- Results from samples of 534 and 853 messages showed 27% and 38% more social network ties detected with the name algorithm than the chain algorithm
- In another set of classes, correlations between these networks ranged from 45% to 69%
- Other work has looked at how the name network is related to perceptions of relationships (as reported by students)
  - For 4 of 6 classes, the name network was consistently more likely to match the self-report network than the chain network

Uso de análise de redes para explorar comportamento no E-learning

- Social Network Patterns
  - Rhythms
  - Participation
  - Networks
- Sample used for initial analysis
  - 8 iterations of the same course
  - 2 per semester Fall 2001 to 2004
- Note
  - The following examples use only the network formed by the **subject line** network

Estatísticas básicas sobre a classe

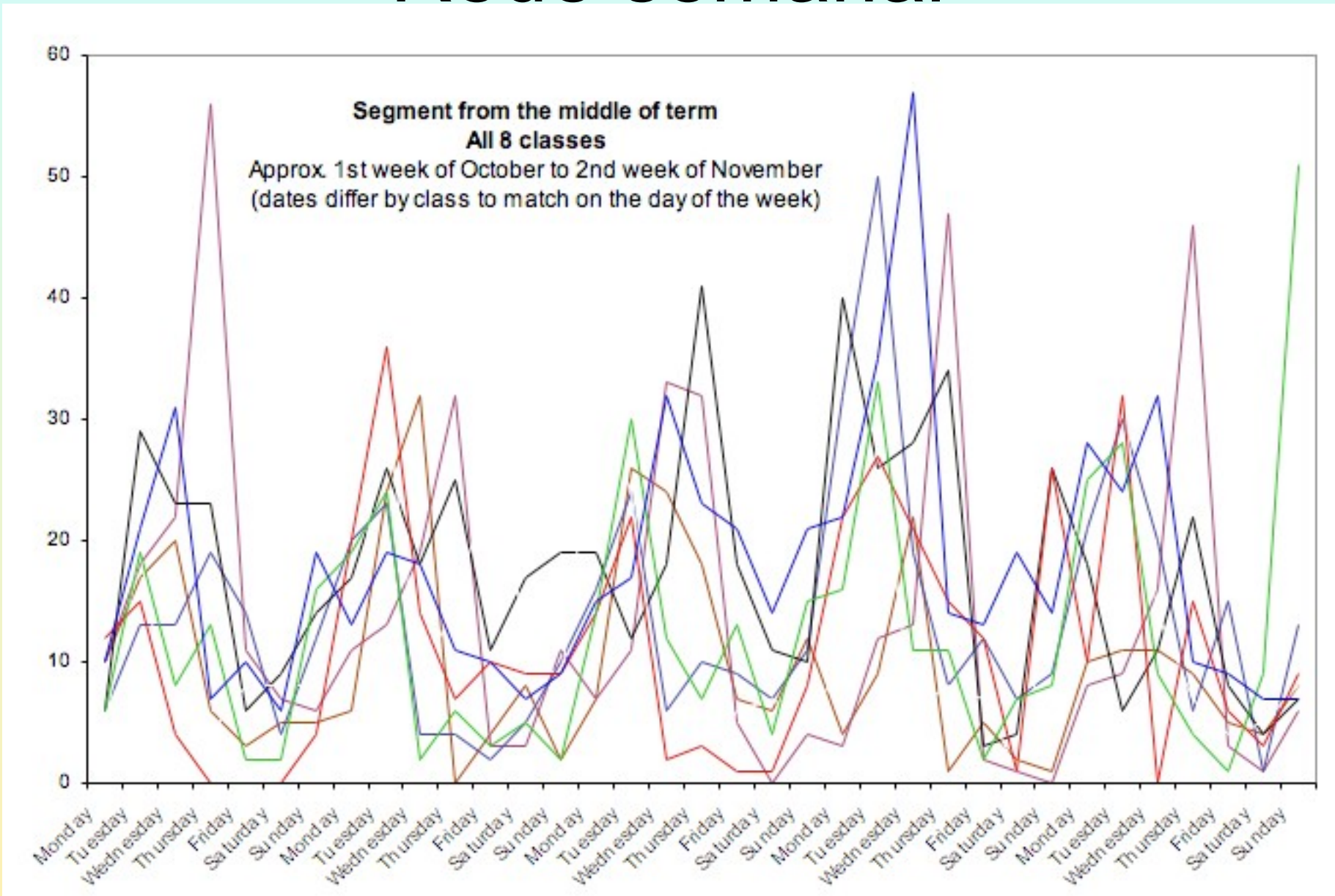
	No. students	No. of instructors / TAs	No. of unique msgs
2001A	33	5	1205
2001B	42	5	1580
2002A	39	4	1469
2002B	46	4	1895
2003A	52	4	1280
2003B	54	4	1242
2004A	31	4	1493
2004B	34	4	2156

- Same course, same instructor each semester
  - Different teaching assistants, adjustments to course content
- Only *public* bulletin boards examined
  - Other communication happened during synchronous lecture sessions, chat, email, private small-group bulletin boards

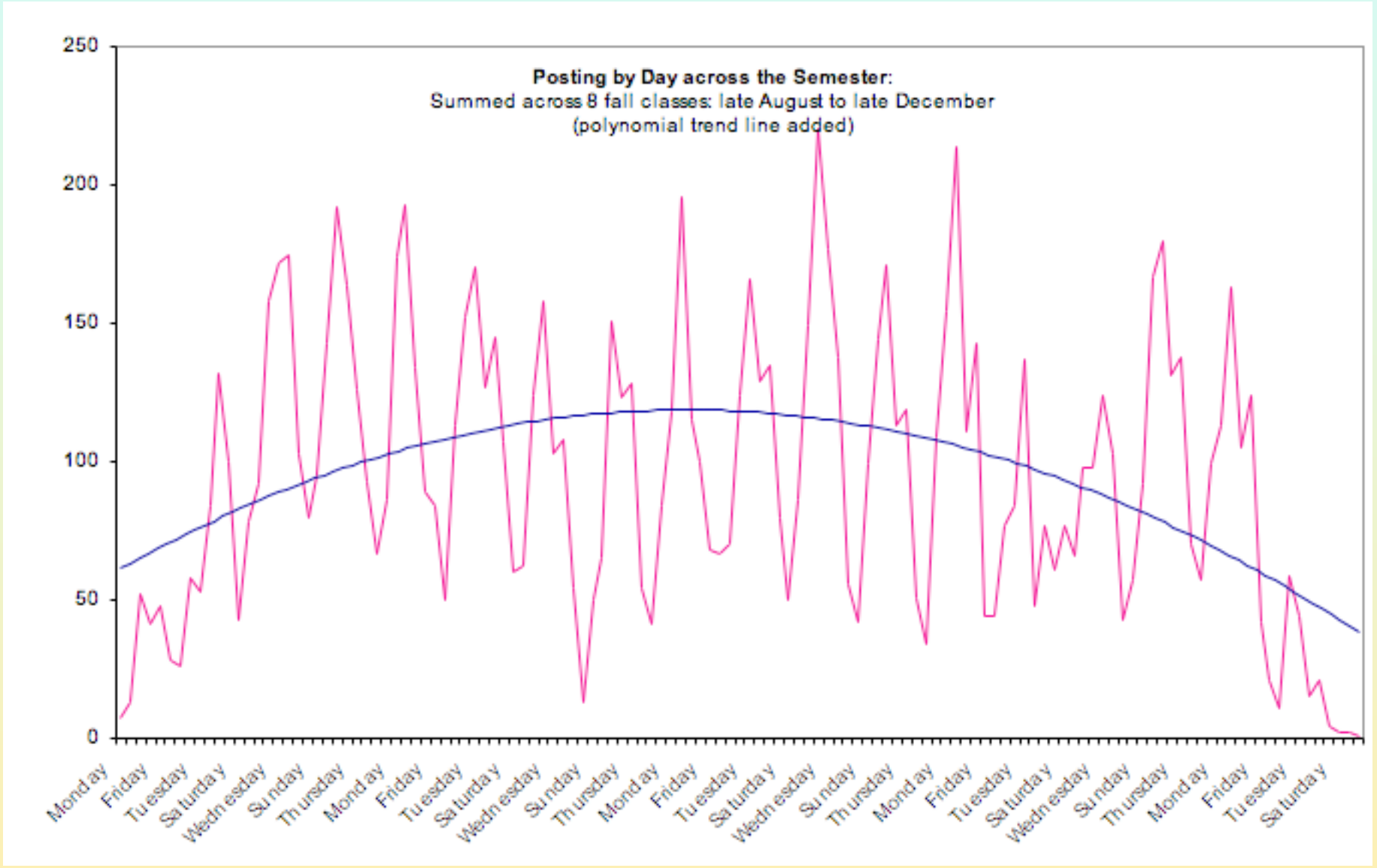
Rítmos

- Existem rítmos nos *postings*?
  - Weekly trend highly evident
  - Semester -- small start-up and finish, but rhythm maintained at approximately the same kind of pace all semester

Rede semanal



Visão do semestre



Comportamentos de *posting*

- Interatividade
  - What is a good interactivity ratio? What is a good response rate density?
    - We don't know
  - Here are some numbers as a baseline from these 8 classes, and as examples of what kinds of numbers might be worth looking at

Interatividade

- Posting Activity
  - Number of participants (Range: 38 - 58)
  - Total number of postings (Range: 1205 - 2156)
- Threading (subject line)
  - Number of threads (373 - 1022)
  - Number of posts per thread (max. 19 - 36)
- Post : Response pairings
  - Direct responding rates (690 - 1144 posts)
  - Direct response network densities (.13 - .42)

Example: Class 2001A

38 participants (33 students + 5)

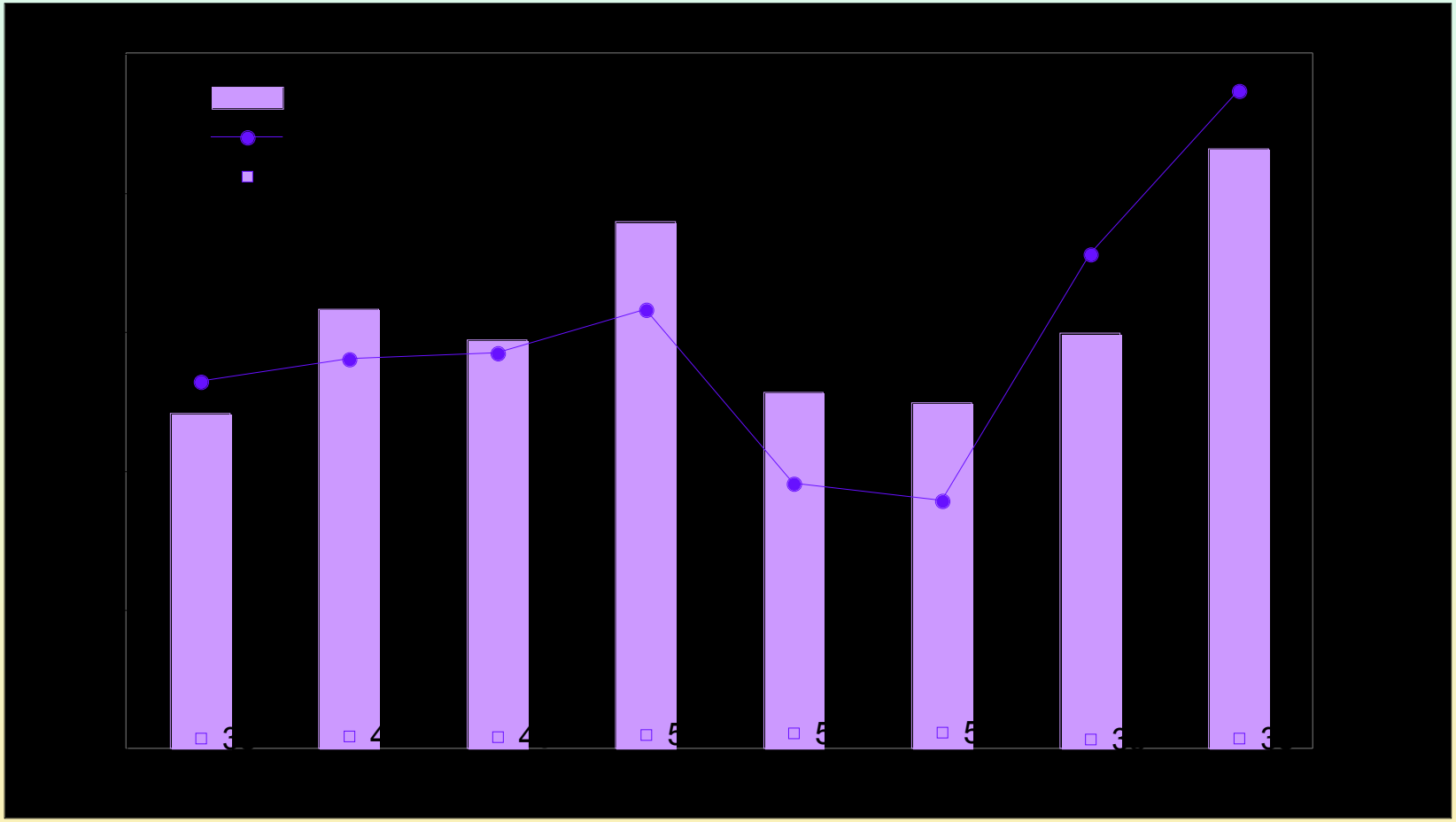
1205 posts in class-wide bulletin boards

373 "threads" as determined by subject line

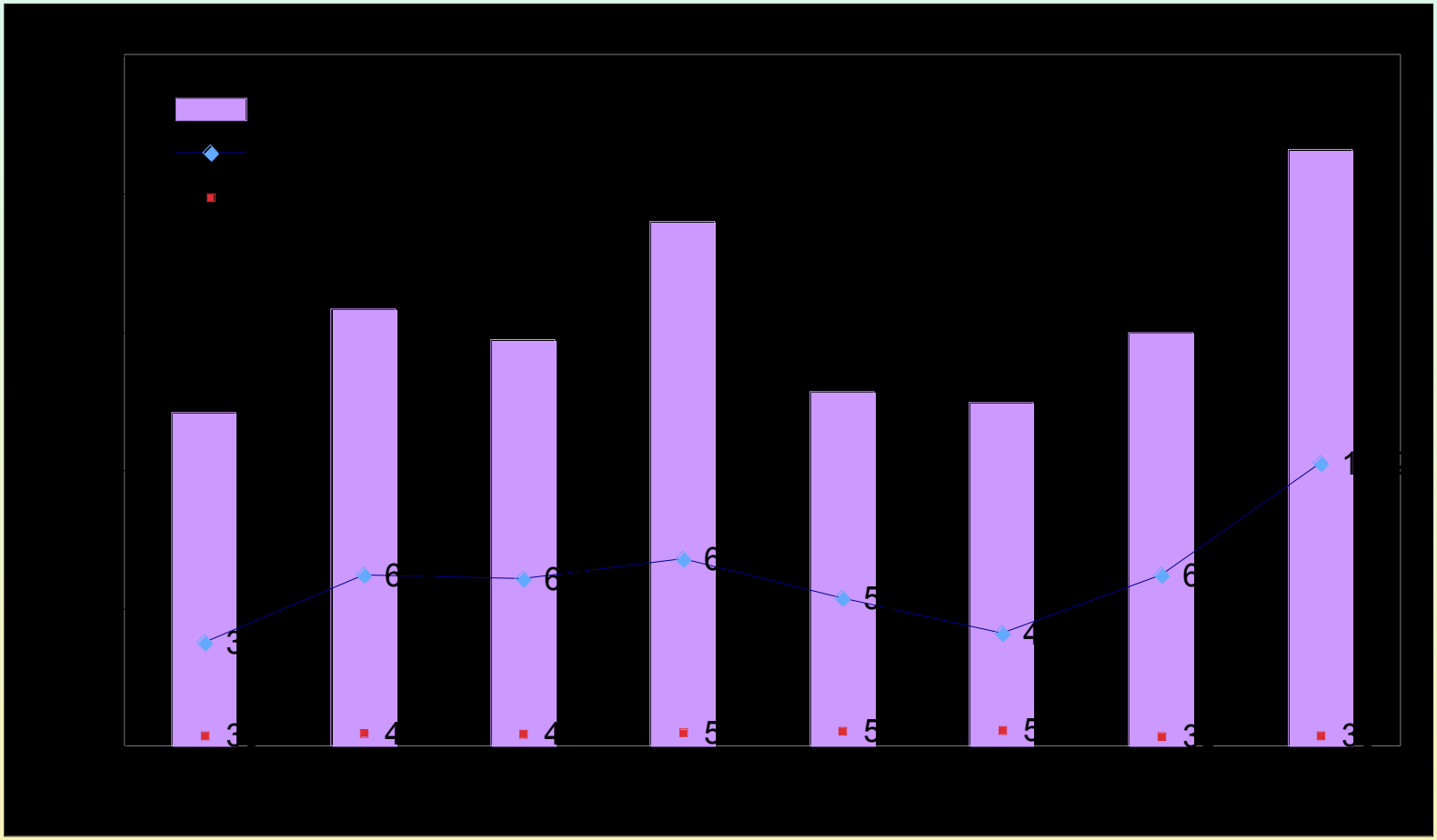
1-24 postings per thread

786 post:responses maintained by **499 pairs**

Atividades de *posting*: 8 classes



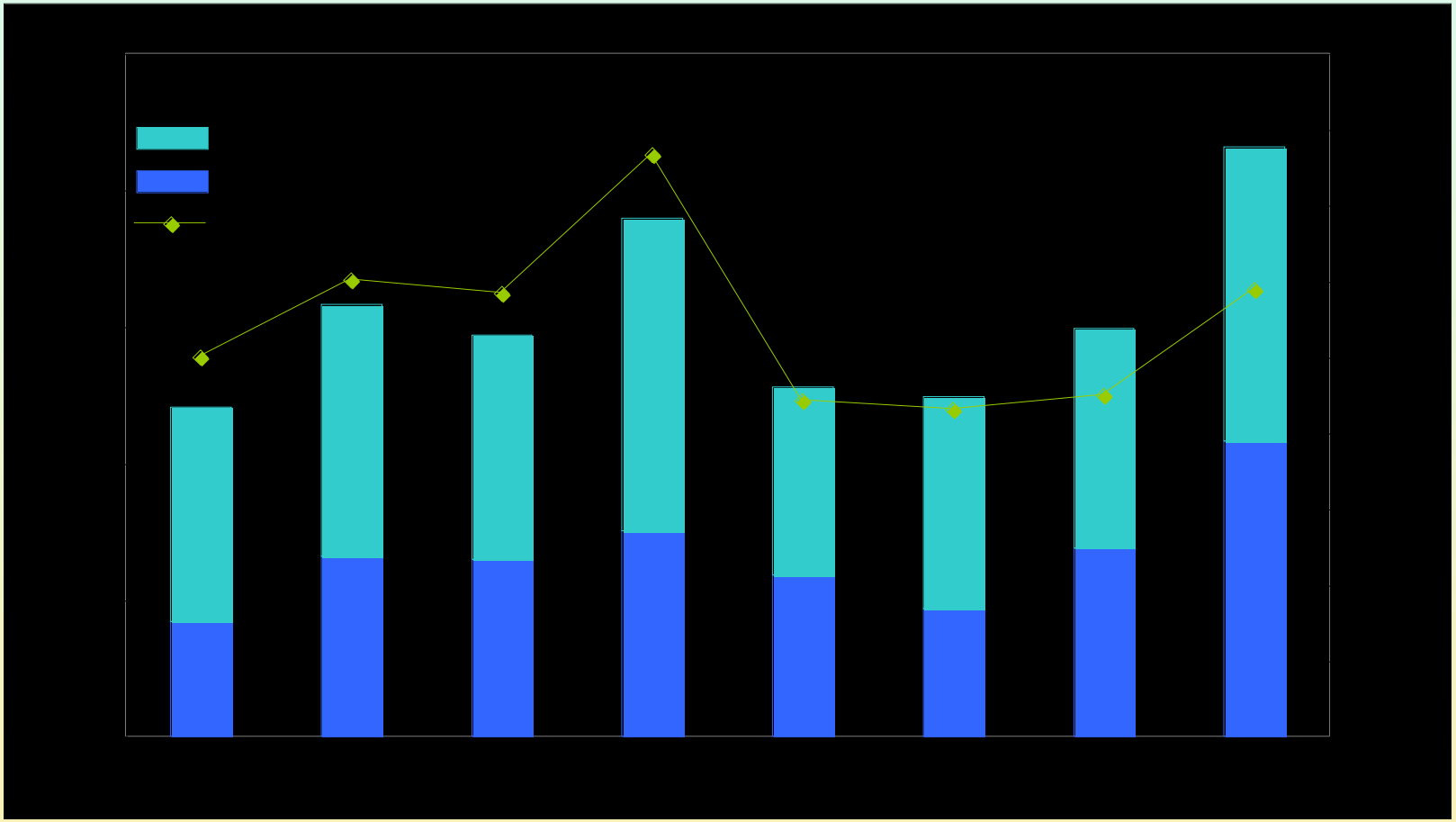
Atividades por tópico: 8 classes



Padrões de conversação

- “Post-Response” Ties
  - A tie is indicated if one participant posts directly after another participant with the same subject line
    - Post from Fred on “Shall we dance?” is directly followed by a post from Ginger with subject line “Shall we dance?”
  - This is a very simple measure of a tie

Padrões de resposta: 8 classes

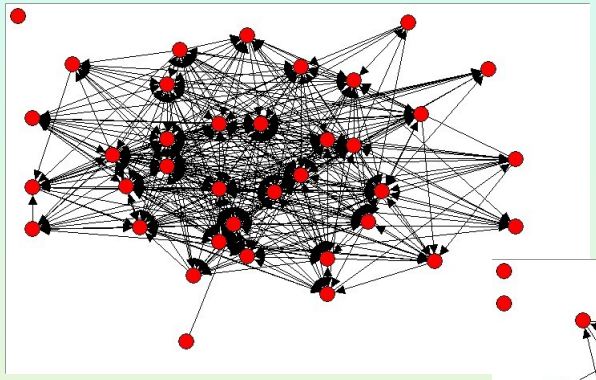


Características das redes

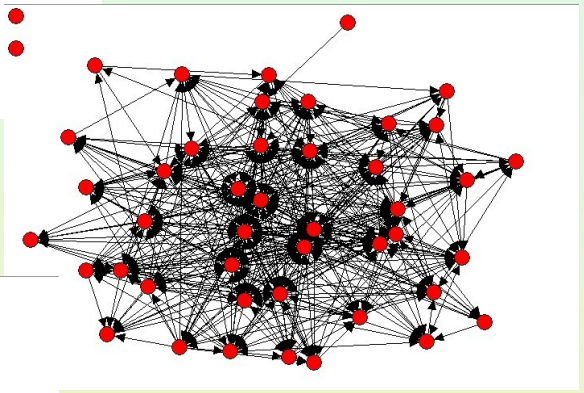
Network density, directed ties

- Tie is considered to be present if a pair has at least one post-response sequence
- Ties are *directed*: ties from A to B are counted separately from ties from B to A
- Density = number of ties / (n x (n-1))

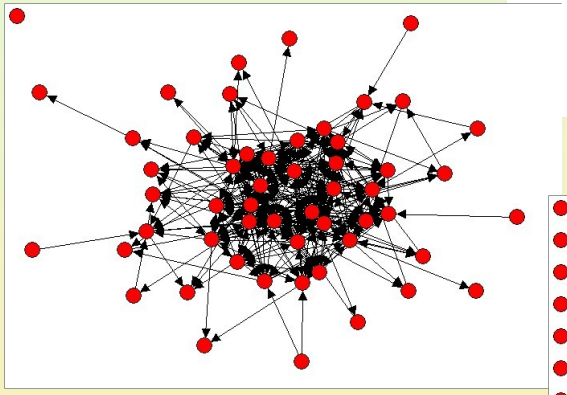
- Example: Class 2001A
  - Number of pairs = 499
  - Number of possible pairs for 38 participants = 38 x 37
  - Density = 499 / 1406 = .35
- Meaning?
  - So far we can only say that they are different
  - Can't say more about 'success' of the class until we check on plans and outcomes

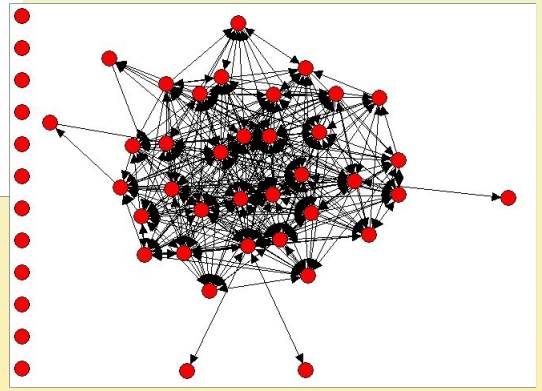


Post-Response tie configurations across 4 classes



Densities (directed)
.35, .32, .14, .38



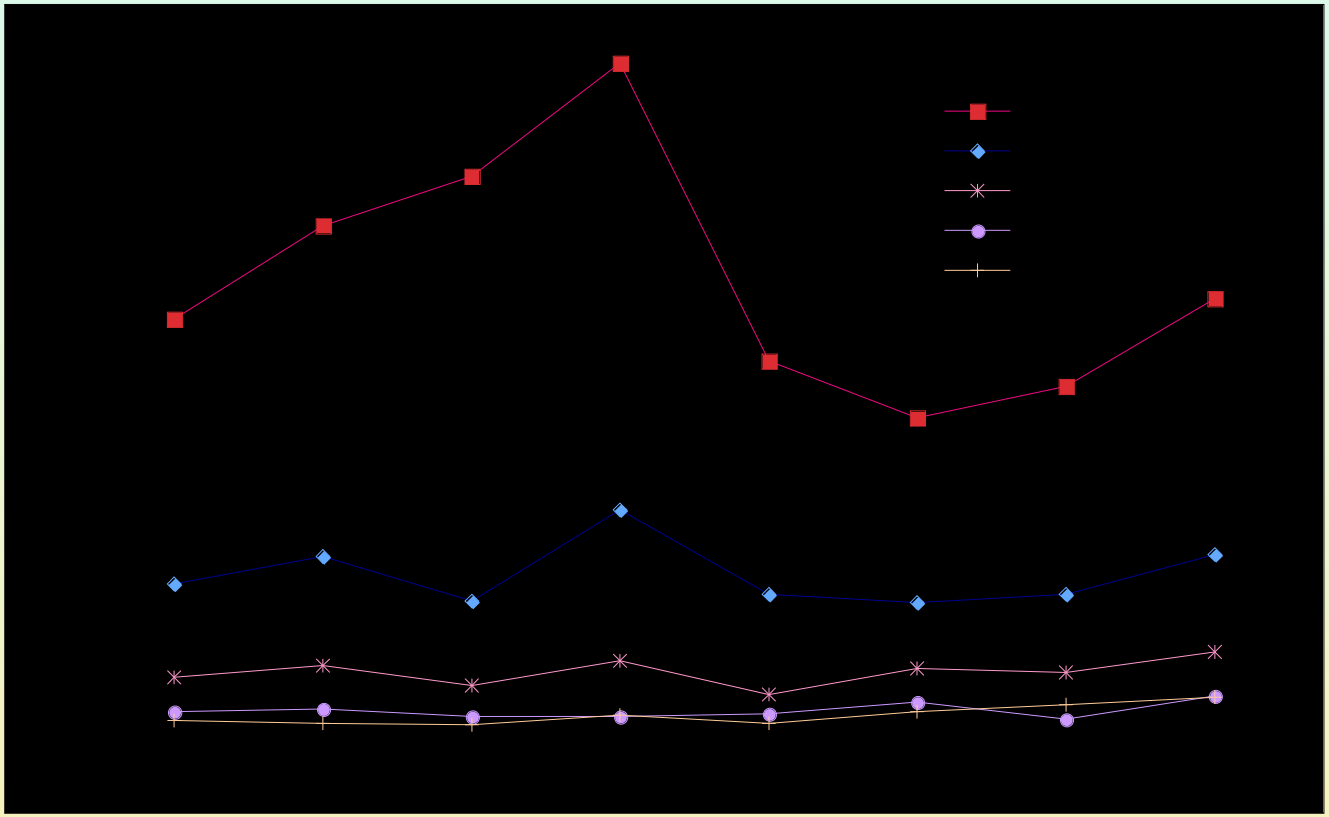


Densidade da interação

	2001A	2001B	2002A	2002B	2003A	2003B	2004A	2004B
No. students + instructors	38	47	43	50	56	58	35	38
No. of possible pairs (n x (n-1))	1406	2162	1806	2450	3080	3306	1190	1406
Number of pairs x number of post : response cases								
1 or more	499	601	583	766	442	430	449	588
2 or more	181	211	155	251	156	188	183	254
Density (directed) x number of post : response cases								
1 or more	.35	.28	.32	.31	.14	.13	.38	.42
2 or more	.13	.10	.09	.10	.05	.06	.15	.18

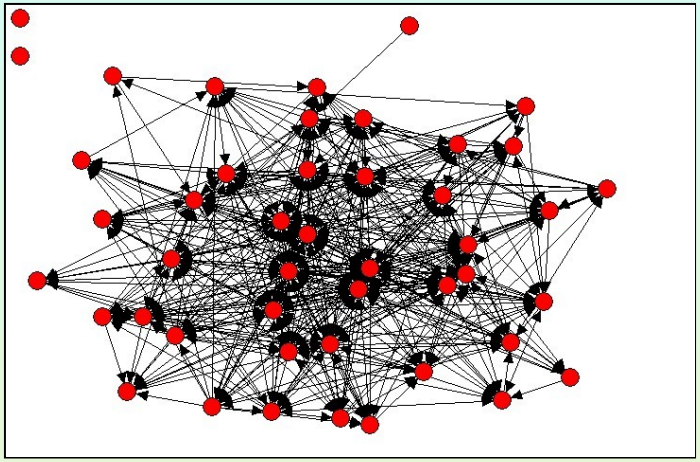
Densities of .13 to .42 for at least one post-response.
Densities fall off rapidly for indication of second post:response (range . 05 to .18). Low densities in particular for 2003A and B.

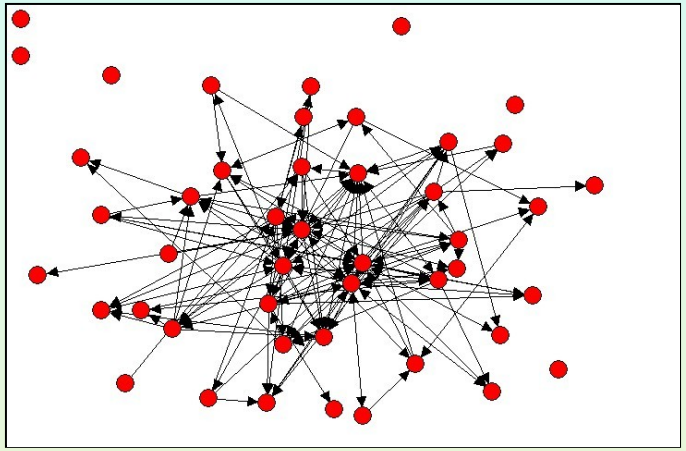
Laços de resposta



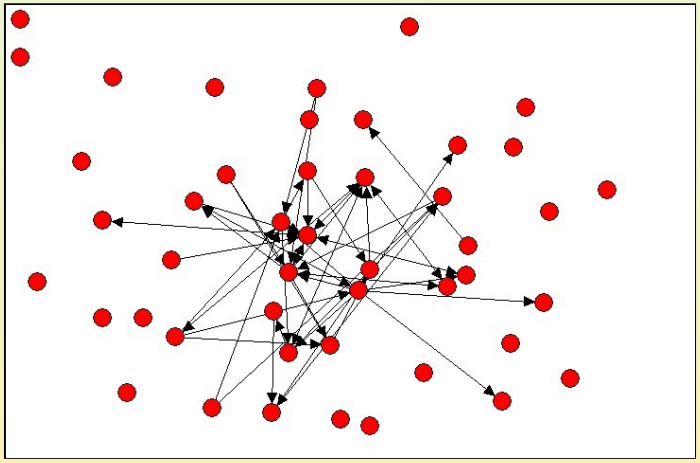
- Most pairs are connected by only **one** immediately following posting (57-73%)
- 17-24% on two subsequent postings; 6-11% on 3; 2-5% on 4; 0-5% on more than 4 iterations
- NB. excludes consideration of multi-way interaction e.g. A<-B, C<-B, A<-C

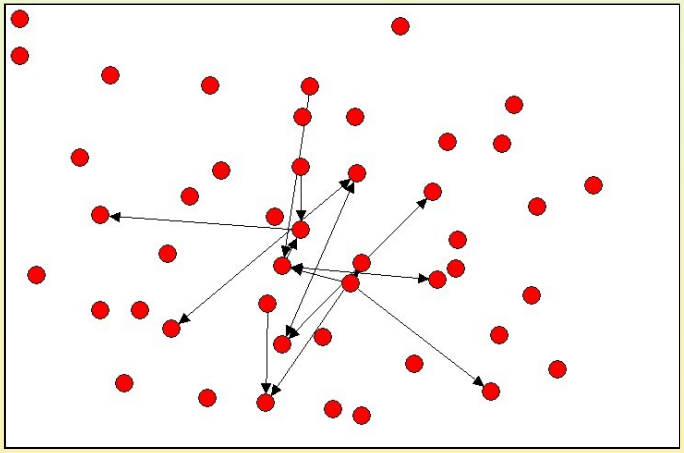
Redes: Densidade, clique, força do laço





Class 2002A: 1 to 4 post:response sequences





Conclusão

- Objetivos e metas
  - Colher e comparar dados de diversas redes
  - Identificar que tipos de interações sociais estão relacionadas a experiências bem sucedidas de indivíduos e grupos
- Até agora
  - Colhendo dados, criando ferramentas automáticas
- Trabalho a fazer
  - Novas análises sobre interações nas classes
  - Análises de outras comunidades online
  - Novas comparações entre dados automáticos e dados fornecidos pelo indivíduo

Referências

- Haythornthwaite, C. & Gruzd, A. (June, 2007). A *noun phrase analysis tool for mining online community*. In C. Steinfield, B.T. Pentland, M. Ackerman & N. Contractor (Eds.). Communities and Technologies 2007: Proceedings of the Third Communities and Technologies Conference, Michigan State University (pp. 67-86). London: Springer.
- Gruzd, A. & Haythornthwaite, C. (2008). *Automated discovery and analysis of social networks from threaded discussions*. International Sunbelt Social Network conference, Jan. 22-27, St. Pete's Beach, Florida.
- Haythornthwaite, C. & Gruzd, A. (2008). *Analyzing networked learning texts*. Paper presented at the Networked Learning Conference, Halkidiki, Greece, May 5-6, 2008. [http:/htl.handle.net/2142/11518]
- Gruzd, A. & Haythornthwaite, C. (forthcoming). Networking online: Cybercommunities. In J. Scott & P. Carrington (Eds.), *Handbook of Social Network Analysis*. London: Sage.